# The Dynamics of Agglomeration Externalities along the Life Cycle of Industries

Neffke, Frank; Svensson Henning, Martin; Boschma, Ron; Lundquist, Karl-Johan; Olander, Lars-Olof

Postprint / Postprint
Zeitschriftenartikel / journal article

**Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:**
www.peerproject.eu

**Empfohlene Zitierung / Suggested Citation:**
Neffke, F., Svensson Henning, M., Boschma, R., Lundquist, K.-J., & Olander, L.-O. (2010). The Dynamics of Agglomeration Externalities along the Life Cycle of Industries. *Regional Studies*, *45*(1), 49-65. https://doi.org/10.1080/00343401003596307

# The Dynamics of Agglomeration Externalities along the Life Cycle of Industries


| | |
|---|---|
| Journal: | *Regional Studies* |
| Manuscript ID: | CRES-2008-0246.R1 |
| Manuscript Type: | Main Section |
| JEL codes: | O18 - Regional, Urban, and Rural Analyses < O1 - Economic Development < O - Economic Development, Technological Change, and Growth, R11 - Regional Economic Activity: Growth, Development, and Changes < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics, R12 - Size and Spatial Distributions of Regional Economic Activity < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics |
| Keywords: | industry life cycle, agglomeration economies, Jacobs' externalities, MAR externalities, evolution, Sweden |

# The Dynamics of Agglomeration Externalities along the Life Cycle of Industries

**Frank Neffke**
**Department of Applied Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: neffke@ese.eur.nl**

**Martin Henning**
**Department of Social and Economic Geography, Lund University, Sölvegatan 12, 223 62 Lund, Sweden. E-mail: martin@keg.lu.se**

**Ron Boschma**
**Section of Economic Geography, Faculty of Geosciences, Utrecht University, P.O. Box 80.115, 3508 TC Utrecht, The Netherlands. Email: r.boschma@geo.uu.nl**

**Karl-Johan Lundquist**
**Department of Social and Economic Geography, Lund University, Sölvegatan 12, 223 62 Lund, Sweden. E-mail: Karl-Johan.Lundquist@keg.lu.se**

**Lars-Olof Olander**
**Department of Social and Economic Geography, Lund University, Sölvegatan 12, 223 62 Lund, Sweden. E-mail: Lars-Olof.Olander@keg.lu.se**



**(Received September 2008: in revised form July 2009)**

ABSTRACT
This paper investigates the changing roles of agglomeration externalities along the industry life cycle. We argue that industries have different agglomeration needs in different stages of their life cycles because their mode of competition, innovation intensity, and learning opportunities change over time. For 12 Swedish manufacturing industries, we determine for each year between 1974 and 2004 whether the industry is in a young, intermediate, or mature stage. Whereas MAR externalities steadily increase with the maturity of industries, effects of local diversity (Jacobs' externalities) are positive for young industries but decline and even become negative for more mature industries.

Keywords: Industry Life Cycle, agglomeration, externalities, evolution, MAR, Jacobs

Classification: O18, R11, R12

1. Introduction

A classical question in economic geography and regional economics is how geographic agglomerations contribute to the economic performance of firms. In the externalities literature, it is commonly understood that diversified cities offer different benefits as compared to specialized cities. A host of empirical studies have focused on the trade-off between diversification and specialization. However, so far, the empirical evidence on agglomeration externalities remains surprisingly inconclusive. The main question we seek to answer is whether the life-cycle stage of an industry determines which type of agglomeration externalities generates the highest benefits for local firms. Although the proposed framework is compatible with theoretical notions like "nursery cities" (DURANTON and PUGA, 2001) and in line with the oft-stated conjecture that externality estimates for high-tech industries differ from those for low-tech industries, to our knowledge, no study has yet systematically investigated the observed divergence in empirical outcomes by linking the concepts of agglomeration externalities to theories of industrial dynamics.

In doing so, we do not treat agglomeration externalities as static, and we will claim that the industry life-cycle framework can greatly help us to understand the changing nature of the effects of the local environment on the efficiency of local industries. Hypotheses are tested with data that cover industry life cycle (ILC) stages for twelve manufacturing industries in a panel of 70 Swedish cities over the period 1974-2004. The proposed statistical framework addresses many of the difficulties that arise in empirical work in the field of agglomeration externalities. Most importantly, we make an effort to reduce the arbitrariness in the choice of spatial units by the use of distance decay functions, and we address the problem of time-invariant regressors in panel data fixed effects models. We start from a production function inspired approach, with value added as the dependent variable. In order to disentangle the effects of factor costs and knowledge spillovers, we use both agglomeration indices and variables capturing differences in factor costs across cities. Our estimates show that young industries benefit from being located in high-cost, high-diversity locations. When moving towards more mature industries, however, the benefits shift gradually to plants located in low-cost, specialized locations.

The structure of the article is as follows. First, in section 2, we briefly discuss the literature on agglomeration externalities. Next, in section 3, we turn to the ILC concept and show how it may be used to structure our expectations about the strength of different types of agglomeration externalities. In section 4, we describe the Swedish dataset and explain how we determine the life-cycle stages of the twelve industries. In section 5, we discuss the econometric specification. Section 6 describes the data and presents the main results of our regression analyses. In section 7, we summarize the main conclusions and research challenges.

## 2. Agglomeration externalities[i]

Agglomeration externalities can be defined loosely as the benefits a firm derives from being located close to other economics actors (see, for an extensive overview, ROSENTHAL and STRANGE, 2004). Often, a distinction is made between three types of externalities: Marshall-Arrow-Romer (MAR) [ii], Jacobs', and urbanization externalities.[iii] These types of externalities can be linked to opportunities for learning and the level of factor costs in a city.

Urbanization externalities are benefits experienced by firms located in large cities. On the one hand, these cities often offer access to large markets (either locally, or because they are hubs in international infrastructure networks), highly educated employees, and strong R&D centres and business services. On the other hand, large cities are high cost environments, with congestion, expensive labour, and high land prices.

MAR externalities arise when firms benefit from a strong local specialisation in their own industry. In the Marshallian tradition (MARSHALL, 1890), MAR externalities can be attributed to three sources: labour market pooling, input-output linkages and intra-industry knowledge spillovers. A large local industry is able to grow and sustain a highly skilled specialised labour force, and has lower matching costs between employers and employees (DURANTON and PUGA, 2004). It will also attract many specialised supplier and customer firms, resulting in lower transport costs and inventories. Moreover, spatial proximity to suppliers and customers facilitates joint innovation efforts along the value chain. Between competitors, knowledge spillovers occur through imitation and skill transfers that are greatly facilitated by face-to-face contacts between geographically proximate actors (STORPER and VENABLES, 2004).

In contrast to MAR externalities, Jacobs' externalities arise when firms benefit from the presence of a high level of industrial diversity. Local diversification gives rise to opportunities for combining knowledge across industries (JACOBS, 1969). Frequently, industries face problems in their production processes that have close analogues in other industries. Solutions that are applied in one industry can, in these cases, often be readily adapted to solve problems in other industries. Moreover, a diversified regional economy increases the likelihood of serendipitous inter-industry knowledge and product combinations to arise. Industrial diversity also results in more stable demand conditions and allows firms to choose from a wide range of local input substitutes, which reduces their exposure to price fluctuations in inputs. A formalization of this notion can be found in the love-of-variety models that use Dixit-Stiglitz (DIXIT and STIGLITZ, 1977) production or consumption functions (see for example DURANTON and PUGA, 2004).

Although congestion and high factor costs in large diversified cities may give rise to negative externalities, as argued above, diversity itself is thought to benefit local industries. However, in some empirical studies, diversity appears to have a negative effect on local economic performance. In fact, DE GROOT et al. (2009) report negative Jacobs' externalities in over half the studies they review. Examples can be found in some of the industries studied by HENDERSON (2003), and in COMBES (2000) for heavy manufacturing industries. Both authors make note of these findings, but they do not probe their meaning or causes. We argue that, in fact, Jacobs' externalities should be expected to become negative if they lead to a lack of focus in the provision of a number of local endowments, such as infrastructure, institutions, local services, and other facilities that are shared by all local industries.

In order to produce efficiently, industries draw on a wide variety of local resources, ranging from professional services to infrastructure and institutions. Although many of these resources are shared across the entire local economy, they can also be customized to a certain degree. For example, business service providers, like legal offices, accountants, and marketing agencies, can specialize their services to clients from a limited number of industries and thereby focus on the specific needs of these clients. Similarly, the literature on regional innovation systems (e.g. COOKE and MORGAN, 1998) has stressed that local (government) institutions and vocational training institutes can become increasingly adapted to the needs of specific local industries. However, the degree to which such local endowments can be customized depends on how many different industries make use of them. For example, although law agencies provide their services to firms in a number of different industries, they will try to gain a specific expertise in the part of the legal code that is relevant to their most important clients. Similarly, the physical infrastructure in a city is shared among all industries, but one can think of many examples where infrastructure was constructed with a particular industry in mind. However, the more heterogeneous the industrial mix of a city, the harder it is to offer tailor-made solutions. Controlling for the size of a city, the degree to which the provision of general (i.e., non industry-specific) local endowments can be adjusted to meet the needs of individual industries depends on the number of different industries that must share these endowments. Therefore, cities with little industrial variety are likely to provide a stronger focus on these general endowments and therefore better meet the specific demands of their local industries. To illustrate this, Figure 1 shows two cities of equal size that can generate just enough demand to sustain two law firms.

- Figure 1 about here -

The law firms are depicted in the lower part of the figure. The upper part of the figure shows the client firms. In city A, there are only four industries, as indicated by the four different shadings for the client firms. Consequently, if we assume that the law firms specialize on clients in specific industries, each of them can focus on the legal code that is relevant for two different industries and still have enough clients to remain in business. By contrast, city B hosts eight different industries. In this city, the size of the local economy is the same as in city A, and we assume total turnover for both law firms to be equal as well. In city B, however, each law firm has to become proficient in the part of the legal code that is specific to four different industries. Therefore, the level of customization offered by the law firms in city B is likely to be lower than in city A. Similarly, services rendered by the infrastructure and institutions in city B may be expected to be less well-adapted to specific local industries. Particularly when an industry's requirements for such local resources do not change too much over time and the local environment therefore has time to adapt to its needs, the lack of such a local focus may represent an important locational disadvantage for such an industry.

In fact, some empirical support for this line of reasoning can be found in an article by COMBES et al. (2004). Combes and his co-authors use two different measures for local diversity simultaneously: a Herfindahl index that assesses how equally employment is distributed across local industries, and a simple count of industries active in a region. They find that "The most favorable local industrial structure would consist in a small number of industries but of roughly the same size …" (p. 237). In other words, local diversity is beneficial to a region, as long as it is not too spread out across too many different industries.

This is fully in line with our claim that a lack of local focus may be harmful for local industries. Although Combes and his colleagues do not elaborate on this finding, this regularity may help explain why some studies find negative Jacobs' externalities.

Table 1 summarizes the different sources of urbanization, MAR, and Jacobs' externalities, differentiating between knowledge spillover effects and the effects of factor costs.

-Table 1 about here-

Since the seminal articles by GLAESER et al. (1992) and HENDERSON et al. (1995), the empirical literature investigating the impact of MAR, Jacobs', and urbanization externalities has expanded rapidly. Unfortunately, though, these efforts have not yet resulted in a complete understanding of agglomeration externalities. FELDMAN's (2000) literature review about the connections between innovation and location notes that there is wide divergence in the empirical results on the importance of MAR economies. GLAESER (2000) reaches similar conclusions when it comes to the difference between the impacts of concentration and diversity, respectively. Also, NEFFKE (2008) finds that the outcomes reported in seven leading articles show considerable differences, even for studies that focus exclusively on the United States. As noted by COMBES et al. 2004, there are many possible explanations for these contradictory findings. For example, studies differ widely in their methodologies: plant-level versus regional studies, panel data versus cross-section analyses, and productivity versus employment regressions. Moreover, samples are drawn from different periods in history and relate to different geographic areas in the world. DE GROOT et al. (2009) reviewed 31 studies containing over 200 parameter estimates. In a meta-analysis, the authors find that both sample issues and methodological issues affect outcomes. However, COMBES (2000), on the basis of a single dataset and a uniform methodology, still finds considerable differences in the externalities that are experienced by different industries. This suggests that another important factor causing divergence of outcomes is industrial heterogeneity.

Referring to a Vernonian product life-cycle view of regional development, it is often hypothesized that product development takes place in large, diversified cities, whereas production takes place in smaller-yet-specialized cities (e.g. HENDERSON, 2003). DURANTON and PUGA (2001) formalize this conjecture in their “nursery cities” concept. We agree that specialized cities may be attractive for different industries than are diversified cities. Indeed, we will argue that this can best be understood by drawing from the literature on industry life cycles.[iv]

3. Industry life cycles and agglomeration externalities

The industry life-cycle framework (ILC) (GORT and KLEPPER, 1982; ABERNATHY and CLARK, 1985; KLEPPER, 1997) is a stylized description of the evolution of an industry from infancy to decline. The archetypical evolution of the output in an industry follows a logistic (or S-) curve, starting with the introduction of a new product, followed by a period of strong expansion of production, which then levels off and eventually leads to a decline. The ILC literature has grown into an extensive body of work with many detailed descriptions and subtleties. In this paper, we restrict the discussion to three aspects of industry life-cycle stages: type of innovation, innovation intensity, and mode of competition. Table 2 summarizes how an industry changes when moving from a young to a more mature stage.

-Table 2 about here-

The birth of a new industry typically follows from radical innovations that result in new products. The young stage is characterised by the development of an immature technology. Innovation intensity is high, as there are many unexplored technological opportunities. Because standardization has yet to set in, large discontinuities in a technological sense are not uncommon. Therefore, as argued by GORT and KLEPPER (1982), information about innovation(s) can come from a wide range of sources in these stages, often from outside the young industry's population of firms. Inter-industry knowledge spillovers associated with Jacobs' externalities should be of utmost importance to these young industries. These industries need – and can accommodate – a large variety of knowledge to build superior products.

Another characteristic of young industries is that firms tend to compete on the basis of quality characteristics of their new products as opposed to their price. Therefore, they are not very sensitive to factor-cost differentials between regions. Given their volatility and age, young industries are usually insufficiently embedded in the local environment to begin to successfully lobby local business service providers and the local government. For this reason, they will be largely unaffected by the lack of focus that may arise in Jacobs' environments. Typically, the scale of non-standardized production in young industries is small. As a consequence, urbanization benefits should not be expected from access to a large market. However, access to a highly qualified labour force and the lead users that can be found in large cities should be beneficial.

.

After a period of experimentation, disruptive technological jumps become less likely and the industry reaches more mature stages. A 'dominant design' (UTTERBACK and SUAREZ 1993) allows for standardization of production. This opens up opportunities to exploit division of labour and economies of scale. Products get more homogenous, output volumes go up, and firms engage increasingly in price competition. This leads to a sharp drop in prices, enlarging the client base from early adopters to a wider general public. The importance of factor costs rises, whereas the mechanization of the production process lowers the demand for highly-educated labour. However, access to large markets and international infrastructure networks is a prime concern in these industries, suggesting a significant effect of urbanization externalities.

As technological opportunities become exhausted, R&D efforts shift towards process innovation in order to increase efficiency. Process innovations often require specialized, industry-specific machines, skills, and knowledge. Such know-how is often of a strong tacit nature and best acquired through processes of the learning-by-doing and imitation that is typical for traditional industrial districts (AMIN, 2003). Inter-industry spillovers become less likely. The value of a local focus increases as the possibilities to tailor the local education system, infrastructure, and many other aspects of the local environment grow (GRABHER, 1993). This may penalize high-diversity environments. Both tendencies lead to lower Jacobs' externalities. Standardization supports the development of a common language and technology framework across firms, facilitating the orchestration of innovation efforts along the value chain. Similarly, industry-specific knowledge can now be carried more easily across firm boundaries by labour mobility. Therefore, opportunities for intra-industry (MAR) spillovers rise.

The ILC description of industry development is highly stylized. In practice, industries could rejuvenate after a radical innovation that has far-reaching consequences for the industry, and casts the industry back into more infant stages. Moreover, at a higher level of industrial aggregation, various technological trajectories are often stacked and overlap with one another. This may obfuscate the life cycle and prevent an industry from progressing through each of the stages in the described order. However, the basic characteristics in each stage are still the same, whether an industry really is new or just rejuvenated. Table 3 merges the elements of tables 1 and 2 to summarize this discussion of the interaction between agglomeration externalities and life cycle dynamics. In the rows, externality types and their different sources are listed. In the columns, we describe how the influence on the competitiveness of local industries changes between negative, neutral, and positive when moving from young to more mature industries.

-Table 3 about here-

4. Data and industry life-cycle stages

Data

The data used in this study cover all Swedish manufacturing plants with five employees or more without further restrictions in the period 1968-1989 and all manufacturing plants with at least five employees that belong to *firms* employing at least 10 people in the period 1990-2004.[v] All plants in the dataset are classified according to the Swedish SNI-code system at the five-digit level (similar to the SIC classification). Due to a change in this system, we had to merge some five digit industries. For these industries, the classes correspond roughly to the three- or four-digit level.[vi] The geographical location of plants is known at the municipality level (there are 277 in our dataset). This provides us with a high-quality database containing detailed information about the development of the Swedish economy from the late 1960s up to 2004.

Calculation of potentials

As the unit of our analysis is the local industry, we aggregate plant-level data into spatial units. A classical problem occurs when simply summing micro-level data up to regional units: outcomes of analyses based on such data are often subject to change when borders of units are redrawn. Moreover, there are many qualitative differences between regional units. For one, some regions are larger in terms of area than are others. Another issue is that some regions are multi-core regions, whereas other regions are dominated by one large city. In this study, we argue that agglomeration externalities are primarily found in cities. Swedish municipalities are, in many cases, too small in terms of population to exhibit the kind of agglomeration externalities discussed in the literature (e.g GLAESER et al., 1992). For these reasons, we construct metropolitan-area data around 70 Swedish labour market regions (so-called functional or A-regions). Most regions are dominated by a single large city, but some harbour two or more equally sized cities. Agglomeration externalities in multi-core regions should be weaker than their combined size suggests, but stronger than what may be expected from each city separately. Simply attributing all economic activity to a single point in space would therefore overstate the size of externalities, whereas focusing only on the largest city would understate them. Instead of committing ourselves to one of these extremes, we calculate potential measures for the largest cities in each of the regions. The lowest spatial level we can distinguish is the municipality. Typically, this consists of a town and some surrounding hamlets or villages. Therefore, it is not too restrictive to assume that all economic activity takes place at the location of the central town. Using a road-distance matrix, we construct a spatially weighted sum of the contributions from all municipalities in Sweden to the largest town in each A-region. This gives us a potential measure for each of the 70 major cities in Sweden. Take as an example the number of plants. Let:

$P_{mit}$ : number of plants in municipality $m \in M$ in industry *i* at year *t*.

$d_{ma}$ : distance by road between the municipality core, *m,* and the A-region core, $a \in A$

*M*: set of all municipalities in Sweden

*A*: set of all A-region cores in Sweden

Now the "plant-potential" in industry *i* for A-region-core $a$ is calculated as follows:

$$P_{ait}^{pot} = \sum_{m \in M} f\left(\delta, d_{ma}\right) P_{mit}$$

$f(\delta, d_{ma})$ is a distance decay function.[vii] Analogously, we can calculate a population-potential, an employment-potential and a valued added-potential for each of the 70 cities. However, if we were to apply the potential calculations as described above to the dependent variable in a regression analysis, we would artificially create spatial autocorrelation. Therefore, when calculating potentials for the dependent variable and the employment variable (which captures the scale of the inputs), we set the contributions of municipalities outside the A-region equal to zero. For all other variables, contributions from all Swedish municipalities are used.

Next, we add data on house prices, acquired from Statistics Sweden.[viii]. The house price indicators, however, are not spatially weighted and reflect the house prices in the 70 core cities.

Industries and lifecycles

In total, our dataset distinguishes between 102 different industries that can be followed consistently over time. Most of the industries are small and are present in only a handful of labour market regions. As this would cause additional estimation problems, we focus on 12 major industries. These 12 account for some 44% of total Swedish manufacturing output in 1974 and for 42% of such output in 2004.

A major challenge is to find an adequate way to determine the life-cycle stages of each industry over time. Traditionally, life-cycle stages have been described using entry-exit patterns (e.g. GORT and KLEPPER, 1982). However, in a small country like Sweden, the number of plants in a single industry and, hence, the number of entrants and exits, is very small.[ix] This means that just a few entries or exits are sufficient to mark the transition between two life-cycle stages. AUDRETSCH and FELDMAN (1996) suggest an alternative approach to identify life-cycle stages. They classify industries into different life cycle stages by looking at the innovation intensity of small firms compared to large firms by drawing on a product announcements database. Young industries are characterized as industries with a high innovation intensity that takes place predominantly in small firms. If large firms undertake the bulk of the innovation effort, industries are said to be either in the growth stage or in the mature stage, depending on whether the overall innovation intensity is high or low. Declining industries are supposed to exhibit little innovation, but that which does exist will be carried out predominantly by small firms. Data on innovation intensity were not available for this study. However, the way we define life-cycle stages has some commonalities with the work by Audretsch and Feldman. Instead of innovation intensity, we look at the market share of young plants in an industry. The logic behind this is that young plants will, in general, use newer production technologies than will mature plants. After all, technology is to a large degree embedded in machinery, which is costly to replace. Moreover, new technologies require new routines and skills, the acquisition of which is costly and time consuming (e.g., NELSON and WINTER, 1982). If an industry is in a stage of strong technological renewal – i.e., if the industry is young or recently rejuvenated – mature plants struggle to respond to a situation where both physical and human capital are becoming obsolete, allowing young plants to capture large shares of the market. In contrast, if the industry is in a stage with a stable technological trajectory, older plants are less threatened by new entrants and will retain a larger share of the market.

Defining old plants as plants that are 10 years or older, we calculate the market share of old plants for all industries in our dataset. To control for changes in the overall plant turnover over the years, we divide the old-plant market shares of each industry by the market share of old plants in the economy as a whole. This yields the over- or underrepresentation of old plants compared to the national level. Next, we normalize this index by subtracting the mean and dividing by the standard deviation across all industries. Let the maturity index be:

$$I_{it} = \frac{VA_{it}^{old} / VA_{it}^{tot}}{VA_{t}^{old} / VA_{t}^{tot}}$$

where:

$VA_{it}^{old}$: value added in old plants in industry *i* at year *t*

$VA_{it}^{tot}$: value added in all plants in industry *i* at year *t*

$VA_{t}^{old}$: value added in all old plants in Sweden at year *t*

$VA_{t}^{tot}$: value added in all plants in Sweden at year *t*

The normalized maturity index is now:

$$I_{it}^{norm} = \frac{I_{it} - mean(I_{it})}{stddev(I_{it})}$$

We distinguish between three levels of maturity: young, intermediate, and mature. In order to obtain a roughly equal number of observations for each type, we use the mean –0.3 times the standard deviation and +0.3 times the standard deviation as demarcation values. Using a five-year moving average to control for business cycle volatility, table 4 shows, for each year, the life-cycle stages for the twelve industries in our study. Industries may move through several stages, such as "other plastics," which progresses through young as well as intermediate and mature stages. However, please note that industries close to the maturity threshold may shift repeatedly between categories.[x] The general picture shows, in line with our intuition, that textiles, sawmilling, carpentry, furniture, paper, and chemicals have been rather mature industries over the entire course of the past three decades. Publishing and communication underwent rejuvenation in the 1990s. Electric motors had entered a young stage already in the 1980s. Other plastics, in contrast, slides into maturity. To a lesser degree, the same holds for metal ware. The instruments industry has been classified as a young industry for all years in our sample.

-Table 4 about here-

5. Econometric specification

In order to measure the size of externalities, we estimate a Cobb-Douglas inspired production function for city-industries. Output is measured by value added. Due to a lack of capital data, the only inputs in the production process at our disposal are employment data. This gives rise to the following multiplicative model:

(1) $$VA_{cit} = T_{cit} L_{cit}^{\alpha} \varepsilon_{cit}$$

where:

$VA_{cit}$ : Value-added potential produced by city c in industry i at year t

$L_{cit}$ : Labour potential employed by city c in industry i at year t

$T_{cit}$ : Technology term. This term contains the externality effects for city c in industry i at year t.

The omission of a variable that captures the size of capital inputs is potentially problematic. For example, it may be particularly expensive to use capital intensive production technology in large cities, where transport and storage of capital goods is costly. In this case, the urbanization externalities parameter would suffer from a downward bias: the negative effect of a lower capital intensity in big cities reduces the parameter estimate for urbanization externalities.

A satisfactory solution would require the estimation of a full Total Factor Productivity (TFP) model as in CINGANO and SCHIVARDI (2004). Such models enable the researcher to quantify the overall efficiency improvements that can be attributed to agglomeration externalities. Without data on capital, this strategy is unavailable. However, concerns about omitted variable bias should be limited as long as the estimate of $\alpha$ stays close to 1. The reason for this is that we can think of labour as a proxy for the total amount of inputs that went into the production process. If we now assume that there are constant returns to scale – which is not unreasonable at the level of regional aggregates – and that there is no confounding with the other variables in the model, we should expect that all effects of the scale of production are captured in $\alpha$. Therefore, a finding of $\alpha = 1$, although strictly speaking not a proof, suggests that labour inputs are also a solid proxy for capital inputs and, as a consequence, that the externality estimates are not confounded by omitted variable bias.

We will now turn to the elements that make up the technology term. Urbanization externalities are measured by the population potential:

$$POPULATION_{ct} = \sum_{m \in M} f\left(\delta, d_{mc}\right) pop_{mt}$$

We would like to disentangle, however, the different aspects of a big city implicit in urbanization externalities. We therefore control for two other factors. The first is the overall wage level in the city. For this purpose, we construct, for each industry, the relative wage level in a city compared to the corresponding national level. We next construct a weighted average, using local employment in each industry as a weight:

$$WAGE_{ct} = \sum_i \frac{L_{cit}}{L_{ct}} \frac{w_{cit}}{\overline{w}_{it}}$$

where:[xi]

$w_{cit}$: average local wage paid in industry *i*.

$\overline{w}_{it}$: average Swedish wage paid in industry *i*.

The resulting index is equal to one if average wages in the city are equal to average wages in the whole of Sweden, keeping constant the composition of industries. As this index is calculated across all 102 manufacturing industries in the economy, the influence of the wage level of a single industry is negligible. The problem of using a wage index is that the wage level in a city reflects both variations in the cost of living across cities and differences in the premium that is paid for the quality of the local workforce. We would like to distinguish between the cost and quality-of-labour effects. As rents account for a large portion of household expenditures, the higher costs of living in big cities are to a large degree accounted for if we control for housing prices ($HOUSE_{ct}$). We will assume that if we keep housing prices constant, the remaining variation in the wage index is due to differences in quality of labour.

MAR externalities have been modelled in various ways in the literature. Levels or shares of own industry employment are widely used indicators (e.g., GLAESER et al., 1992, HENDERSON et al., 1995, HENDERSON, 1997). However, these indicators do not differentiate between plant-internal and plant-external economies of scale. In the extreme case, where all the employment in a city-industry is located in a single plant, the effects of local employment are fully attributable to internal economies of scale. The number of plants, in contrast, can only give rise to external economies of scale.[xii] HENDERSON (2003) argues that the number of plants is a good measure for capturing MAR externalities. His reasoning is that each plant can be interpreted as an experiment with a specific variation on the industry's production process. The potential of intra-industry knowledge spillovers depends, therefore, on the local number of those experiments, not so much on the number of the local industry's employees. Moreover, workers acquire more industry-specific skills if they can become employed and educated in different local firms. The number of plants also measures the number of potential innovation partners in the own industry. For these reasons, we measure the scope for knowledge spillovers in MAR externalities as the number-of-plants potential in the local industry:

$$MAR_{cit} = \sum_{m \in M} f(\delta, d_{mc}) P_{mit}$$

where:

$P_{mit}$ : the number of plants in municipality *m*, industry *i* at year *t*.

Jacobs' externalities result from the presence of a large diversity of industrial activity in a city. To quantify this, many authors use an index that measures the evenness of the distribution of economic activity across different local industries, like the Hirschman-Herfindahl index (HHI) or the entropy index of local employment shares. However, these measures are not very informative, as constellations with very different spillover potentials may yield the same HHI.[xiii] The number of significant industries in a city, by comparison, is a more adequate measure. We call an industry's presence in a region significant if its size reaches a certain threshold. As we argued that spillovers are related to the number of experiments in different plants, we base the Jacobs' externalities indicator on the number of plants potential.[xiv]

$$JACOBS_{ct} = \sum_{i} g\left( \sum_{m \in M} f(\delta, d_{mc}) P_{mit} \geq 10 \right)$$

where:

$g(.)$: indicator function that evaluates to 1 if its argument is true and 0 otherwise.

Assuming that all externalities and control variables enter the technology term in a multiplicative way, we arrive at the following log-transformed estimation equation:

(2)
$$\log(VA_{cit}) = cst + \alpha \log(L_{cit}) + \beta_1 \log(POPULATION_{ct-2}) + \beta_2 \log(WAGE_{ct}) +$$
$$\beta_3 \log(HOUSE_{ct}) + \beta_4 \log(MAR_{cit-2}) + \beta_5 \log(JACOBS_{ct-2}) + \log(\varepsilon_{cit})$$

The effects of local learning will only be felt after a certain amount of time. We therefore use a two-year lag for all variables that mainly capture knowledge spillovers.[xv] To meet the aim of the study, we must discover how externalities differ across ILC stages. We therefore pool observations across all industries and make coefficients dependent on the particular ILC stage. Taking the lag 2 into consideration, we end up with a panel of 70*12 city-industries for 34 years. Adding city-industry fixed effects and year dummies, we get:

(3)
$$\log(VA_{cit}) = cst + \alpha^s \log(L_{cit}) + \beta_1^s \log(POPULATION_{ct-2}) + \beta_2^s \log(WAGE_{ct}) +$$
$$\beta_3^s \log(HOUSE_{ct}) + \beta_4^s \log(MAR_{cit-2}) + \beta_5^s \log(JACOBS_{ct-2}) + \eta_{ci} + \delta_t + \log(\varepsilon_{cit})$$

where:

$\eta_{ci}$: city-industry fixed effects

$\delta_t$: year fixed effects

$.^{s}$ : superscript indicating industry life-cycle stage: young / intermediate / mature

## 6. Empirical results

Table 5 shows the main descriptive statistics for our dependent variable, log(VA), and for the six regressors. Table 6 contains the correlations between the regressors.

-Table 5 about here-

-Table 6 about here-

As a baseline, we estimate the effects of externalities without distinguishing between life-cycle stages using a fixed effects (FE) estimation. The outcomes are shown in column (1) of table 7. The estimate for employment (L) is very close to 1, indicating constant returns to scale. In fact, this finding is replicated in all the estimations of this article. As argued before, this is an indication that the omitted variable bias due to a lack of data on capital inputs is limited. MAR externalities are positive and have an elasticity of about 2%, meaning that a doubling of the number of own-industry plants leads to efficiency gains of 2%. As a comparison, this estimate ranges from 2% to 8% in HENDERSON (2003). Jacobs' externalities seem to be absent. The point estimate for urbanization externalities, as measured by population, is high, but the standard error is large as well. The reason for this is apparent in table 5. The population of a city changes only very slowly over time, which gives rise to a very low within standard deviation. However, the between standard deviation is high, as, cross-sectionally, population potentials vary widely. In FE models, this leads to imprecise estimates. Including the extra information on wages and housing prices (column 2, table 7) unfortunately does not solve the problem.

One way to improve the precision of the estimates is using random effects (RE). However, RE models assume that the unobserved city-industry fixed effects are uncorrelated with the regressors. Theoretically, there is little reason for this assumption. Column 3 (table 7) presents the results. The population estimate is small but negative, and the standard error drops tremendously. The other parameter estimates have remained more or less the same. However, a Hausman test on the adequacy of the RE model rejects the RE specification at any conventional level.

This leaves us with a dilemma: on the one hand, FE will not allow us to get precise estimates on one of our core variables; on the other hand, RE has not passed the Hausman test. Theoretically, the Hausman-Tailor procedure (HAUSMAN and TAYLOR, 1981) could be applied. However, the fact that, for this method, one must find variables that can be convincingly thought of as a priori uncorrelated with the city-industry effects makes this method problematic (ARELLANO, 2003, p. 44).

A different solution has been developed by PLÜMPER and TROEGER (2007). These authors use a procedure originally proposed by HSIAO (2003).[xvi] In their "Fixed Effects Vector Decomposition" (FEVD) method, variables can be modelled as either time-varying or (predominantly) time-invariant. The time-varying variables can be estimated without bias. However, the estimates of the effect of time-invariant and slowly changing variables are biased, as already noted by HSIAO (2003). This bias can be interpreted as a between-groups (cross-sectional) omitted variable bias and therefore depends on the correlation between the time-invariant variables and the unexplained city-industry effect. As the city-industry effects are unobserved, it is impossible to assess this correlation. However, Plümper and Troeger show that, for a wide range of values, in small samples, the FEVD estimator outperforms the FE, RE, and Hausman-Taylor estimators in terms of the Root Mean Squared Error (RMSE). In other words, the increased efficiency more than offsets the modest bias in our estimates.

In our study, the superiority of the FEVD depends on two variables: first, the correlation between the city-industry effects and the regressors that we regard as slowly changing and, second, the ratio of between to within standard deviations of the slowly changing variables. Looking at tables 5 and 6, population is clearly a slowly changing variable, with a between-to-within ratio of over 17. As POPULATION is strongly correlated with JACOBS, and a large part of the variation in JACOBS is actually cross-sectional, omitting JACOBS from the residuals would give rise to a large omitted variable bias in the second step of the FEVD procedure.

Column (4) in table 7 shows the estimates for the FEVD specification, where POPULATION and JACOBS are modelled as slowly changing variables. The estimates are similar to the RE estimates. The parameter estimates for the time-varying variables, L, MAR, HOUSE, and WAGE, are all indistinguishable from their FE estimates in column (2). However, the standard errors of FEVD estimates are considerably lower. Jacobs' externalities are significant and negative. The point estimate for population is positive and, for the first time, significant (though very close to the RE estimate). Finally, the estimate for WAGE is positive and significant.

The ILC and agglomeration externalities

Having arrived at a satisfactory econometric specification, we now turn to our research question: how do externalities change with life-cycle stages. Column (5) of table 7 shows the outcomes of the final model. Parameters are allowed to vary depending on the life-cycle stage of the industry. Economies of scale for industries in all stages are more or less constant, as indicated by the employment parameter estimates that are very close to 1. MAR externalities are positive in all industries. The elasticity estimates, however, clearly rise from hardly significant (0.8%) in young industries to 1.3% in intermediate industries and to 2.2% in mature industries. The pattern for Jacobs' externalities runs opposite to this. Young industries benefit from local diversity (doubling diversity leads to a rise in efficiency of 1.9%), whereas mature industries experience large negative effects (-5.3%). Both estimates are highly significant. The estimate for intermediate industries is small and insignificant. These outcomes support our hypothesis that young industries, with their low levels of standardization, are open to knowledge from very diverse sources, but do not benefit much from specialized, industry-specific knowledge. Mature industries, on the other hand, benefit far more from intra-industry knowledge spillovers, but experience difficulties in diversified cities, in line with our hypothesis that these industries would benefit from a focused local environment.

As explained in section 2, the size of the local population can benefit or harm a local industry in various ways. The higher costs of living and the higher quality and level of education in large cities are controlled for by the variables WAGE and HOUSE. Even so, access to large and sophisticated markets gives rise to positive effects, whereas congestion has negative consequences. According to our estimates, the net benefit derived from locating in large cities is positive for mature and intermediate industries, but negative for young industries. This is surprising, but suggests that big city amenities and immediate market access only play a role for intermediate and mature industries.

Young industries need a highly educated labour force to cope with the volatile nature of this part of the technological trajectory. In mature industries, production processes have become more standardized and mechanized. For these, the quality of labour plays a smaller role. This is reflected in the WAGE estimates. The benefits of high wages are indeed limited to young industries. Column (6) shows outcomes when housing prices are omitted. The positive effect for young industries has decreased sharply, supporting the claim that housing prices effectively control for the factor-cost component in the WAGE variable.

High housing prices should harm all industries alike. For mature and intermediate industries, this is confirmed, although more so for intermediate than for mature industries. Surprisingly, however, young industries benefit from higher housing prices. One reason for this could be that housing prices are, to some extent, correlated with factors affecting the quality of life in a city. Given the negative effect of POPULATION in young industries, this suggests young industries thrive in smaller but highly-developed, expensive cities.

Robustness: endogeneity

A potential problem in the interpretation of our outcomes is the presence of endogeneity. Highly productive cities may attract people and firms just as much as a high number of local plants or inhabitants may cause local industries to be more productive. Moreover, wages and housing prices may rise because of productivity increases in local industries. The statistical association between VA and these variables may therefore also be a result of a reverse causality. Such an argument builds largely on feedback loops through effective local demand.

The standard econometric approach to accommodate such problems is instrumental variable estimation. For this technique to work, instruments are needed that are correlated with the regressors, yet uncorrelated with VA. We do not know of any variables that qualify for this task.[xvii] In panel data, the time-series dimension often provides the researcher with suitable instruments, which can be used in General Method of Moments (GMM) estimation. In this approach, first-differenced equations are estimated using lagged independent variables as instruments. The applicability of this method critically depends on the strength of the variables (i.e., on the correlation between lagged regressors and first-differenced regressors). In our case, experiments with GMM led to highly unstable coefficients. This is not surprising, as, for our independent variables, the variation in lags explained typically about as little as 5% of the variation in first-differences. Such values suggest that the GMM approach is indeed unlikely to yield any interpretable results. Instead of IV estimation, we therefore tried to make sure that reverse causality is likely to be an order of magnitude smaller than modelled causality.

First, we measure the dependent variable at the city-industry level, whereas most regressors - population, house prices, diversity, and wages - are measured at the city level. The 12 industries in the study together make up under 45% of all manufacturing employment. Manufacturing as a whole accounts for roughly one third of Swedish employment. Therefore, the variation in local demand as a consequence of the variation in VA in one of the investigated industries – the feedback loop causing reverse causality – will be negligible when compared to the modelled causality.

Second, for our externality indicators, we use a lag-2 specification. Take, for example, the MAR externalities.[xviii] MAR is measured two years before high productivity is manifest. Assuming that firms are unable to anticipate high productivity, but rather react to it, new plants should become operational after the high productivity is realized, not before. Again, this should weaken reverse causality.

The third reason we are not too concerned about endogeneity is that, in some cases, our estimates should be regarded as cautious estimates because we would expect reverse causality to run in the opposite direction of modelled causality. Housing prices, for example, should rise as a reaction to an increase in VA. For two out of three industry stages, however, we find a negative relation between housing prices and VA. The same argument holds for the influence of wages in mature and intermediate stages.

Finally, we are not interested in the level of the parameter estimates per se, but rather in their changes across industry life-cycle stages. There is no compelling reason to expect that, for instance, housing prices will react differently if the VA increase is generated in a young industry or in a mature industry. The increase in local demand as a consequence of a local

VA surge should have the same impact on the regressors regardless of the industry in which the VA is produced. Moreover, as we will show below, if we move between different cut-offs for our ILC stage definitions, coefficients behave exactly as implied by our theoretical framework. Any effective reverse-causality charge to our findings must also have an adequate explanation for this observation. Taking all this together, we deem it unlikely that endogeneity invalidates the main results of this article.



Robustness: changing period definition

The results above obviously depend on the definition of the industry stages. Therefore, we rerun model (5) in table 7 for two alternative life-cycle stage definitions. Column (5a) in table 8 shows the results when the category of intermediate industries is narrowed and the line demarcating mature and young industries gets thinner. Column (5b) uses the same limits as in table 7, and, in column (5c), the intermediate category is widened at the expense of the young and the mature categories. As we move from (5a) to (5b) to (5c), industries classified as mature become increasingly older, and industries classified as young become increasingly younger. If our hypotheses are correct, the parameter estimates for young and mature industries should lie closer together in column (5a) and further apart in column (5c).

Indeed, the patterns for both MAR and Jacobs' externalities are amplified when moving from column (5a) to (5b) and then to (5c). Also, the variables that measure urbanization externalities get more pronounced. For all variables – POPULATION, HOUSE, and WAGE, - the difference between the estimates for mature and young industries increases. This strongly supports the robustness of our analyses.

-Table 8 about here-

7. Conclusions

In this paper, we conjectured that industries have different agglomeration needs in different phases of their life cycles. To test this, we set up a framework that describes the evolution of agglomeration externalities along the life cycle of industries.

In all, our results show that the benefits industries derive from their local environment are strongly associated with their stage in the industry life cycle. Moving from young to intermediate to mature industries, the benefits derived from local specialization steadily increase. In contrast, the benefits from local diversity for young industries are positive, then turn insignificant for intermediate industries, and finally become negative for mature industries. These findings support our hypothesis that with increasing levels of maturity, industries experience rising benefits of intra-industry spillovers, but declining inter-industry spillovers. Moreover, we have argued that the relative stability of mature industries would allow them to take advantage of more specialized environments and make them more vulnerable to a lack of local focus that is more common in diversified cities. This could explain the negative Jacobs' externalities we found in these industries. We also show that, in line with our ILC framework, factor costs and quality of labour have very different effects on the efficiency of young and mature industries. Whereas the former thrive in expensive, medium-sized cities with highly qualified and costly labour, the latter are better off in low-cost cities with a relatively large local market.

Instead of treating agglomeration externalities as static, our results show that the study of agglomeration externalities demands a dynamic, long-term perspective. The industry life-cycle framework can greatly help us understand the changing nature of effects of the environment on the efficiency of local industries.

In terms of methodology, we used a procedure that solves some of the main issues encountered in empirical research in this field. For instance, the use of city-potentials smoothes regional borders and reduces arbitrariness in the choice of regional units. Another improvement was the FEVD methodology to estimate the impact of population, which typically shows very little time-variation. Finally, we disentangled many of the different aspects of urbanization externalities by using information on local housing prices and wages.

There are, however, still some open questions. Although we have argued that endogeneity issues are of no great concern, an analysis at the plant level could solve the issue in a more theoretically satisfactory way. Moreover, replication of our results in other countries and industries, especially in service industries, is necessary. Finally, juxtaposing diversity and specialization unnecessarily dichotomizes the local economic environment into a similar, MAR-externalities-generating component and a dissimilar, Jacobs'-externalities-generating component. In fact, as argued by FRENKEN et al. (2007): "Analogous to economies of scope at the firm level, one expects knowledge spillovers within the region to occur primarily among related sectors, and only to a limited extent among unrelated sectors." (p. 688). By implication, quantifying the relatedness linkages between industries in the economy, as for instance in NEFFKE and SVENSSON HENNING (2008), is necessary to arrive at a more complete understanding of the agglomeration effects that occur as a consequence of a specific set of industries at the regional level.

References

ABERNATHY W. J. and CLARK K. B. (1985) Innovation: Mapping the Winds of Creative Destruction, Research Policy 14, 3-22.

AMIN A. (2003) Industrial districts, in SHEPPARD E. and BARNES T. J.(Eds) A Companion to Economic Geography, pp. 149. Blackwell, Oxford.

ARELLANO M. (2003) Panel Data Econometrics. Oxford University Press, Oxford.

AUDRETSCH D. B. and FELDMAN T. J. (1996) Innovative Clusters and the Industry Life Cycle, Review of Industrial Organization 11, 253-273.

BOSCHMA R. A. (1997) New Industries and Windows of Locational Opportunity. A Long-Term Analysis of Belgium, Erdkunde 51, 12-22.

CINGANO F., and SCHIVARDI F. (2004) Identifying the Sources of Local Productivity Growth. Journal of the European Economic Association 2(4), 720–742.

COMBES P-P. (2000) Economic Structure and Local Growth: France, 1984-1993, Journal of Urban Economics 47, 329-355.

COMBES P.-P., MAGNAC T. and ROBIN J.-M. (2004) The dynamics of local employment in France, Journal of Urban Economics 56, 217-243.

COOKE P.H. and MORGAN K. (1998) The Associational Economy. Firms, Regions, and Innovation. Oxford University Press, Oxford.

DE GROOT H., POOT J. and SMIT M. J. (2009) Agglomeration Externalities, Innovation and Regional Growth: Theoretical Perspectives and Meta-Analysis, in CAPELLO R. and NIJKAMP P. (Eds) Handbook of Regional Growth and Development Theories, pp. 256-281. Edward Elgar, Cheltenham.

DIXIT A. K. and STIGLITZ J. E. (1977) Monopolistic competition and optimum product diversity, The American Economic Review 67 (3), 297-308.

DURANTON G. and PUGA D. (2001) Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products, The American Economic Review 91 (5), 1454-1477.

DURANTON G. and PUGA D. (2004). Micro-foundation of urban agglomeration economies, in HENDERSON J. and THISSE J.-F. (Eds) Handbook of Urban and Regional Economics, vol 4, pp. 2065-2118. Elsevier, Amsterdam.

FELDMAN M. P. (2000) Location and Innovation: the New economic Geography of Innovation, Spillovers, and Agglomeration, in CLARK G. L., FELDMAN M. P. and GERTLER M. S. (Eds) The Oxford Handbook of Economic Geography, pp. 373-394. Oxford University Press, Oxford.

FRENKEN K., VAN OORT F.G. and VERBURG T. (2007) Related Variety, Unrelated Variety and Regional Economic Growth, Regional Studies 41 (5), 685-697.

GLAESER E. L. (2000) The new economics of urban and regional growth, in CLARK G. L., FELDMAN M. P. and GERTLER M. S. (Eds) The Oxford Handbook of Economic Geography, pp. 83-98. Oxford University Press, Oxford.

GLAESER E. L., KALLAL H. D., SCHEINKMAN J. A. and SCHLEIFER A. (1992) Growth in Cities, The Journal of Political Economy 100 (6), 1126-1152.

GORT M. and KLEPPER S. (1982) Time Paths in the Diffusion of Product Innovations, The Economic Journal 92 (367), 630-653.

GRABHER G. (1993) The Weakness of Strong Ties: the Lock-in of Regional Development in the Ruhr Area, in GRABHER G. (Ed) The Embedded Firm. On the Socioeconomics of Interfirm Relations, pp. 255-277. Routledge, London and New York.

HAUSMAN J. A. and TAYLOR W. E. (1981) Panel data and unobservable individual effects, Econometrica 49, 1377-1398.

HENDERSON J. V. (1997) Externalities and industrial development, Journal of Urban Economics 42, 449-470.

HENDERSON J. V. (2003), Marshall´s Scale Economies. Journal of Urban Economics 53. 1-28.

HENDERSON J. V., KUNCORO V. A. and TURNER M. (1995) Industrial Development in Cities, Journal of Political Economy 103 (5), 1067-1090.

HSIAO C. (2003) Analysis of panel data. Second edition. Cambridge University Press, Cambridge,.

JACOBS J. (1969) The Economy of Cities. Vintage, New York.

KLEPPER S. (1997) Industry Life Cycles, Industrial and Corporate Change 6 (1), 145-181.

MARSHALL A. (1890) Principles of Economics. (1920 edition). Macmillan, London.

NEFFKE F. M. H. (2008) Time varying agglomeration externalities: A long-term perspective on the changing benefits of agglomeration in UK Counties (1841-1971), PEEG Working Paper Series #08.18.

NEFFKE F. M. H. and SVENSSON HENNING M. (2008) Revealed Relatedness: Mapping Industry Space, PEEG Working Paper Series #08.19.

NELSON R. R. and WINTER S. G. (1982) An Evolutionary Theory of Economic Change. Bellknap Press/ Harvard University Press, Cambridge, Mass..

PLÜMPER TH. and TROEGER V. E. (2007). Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects, Political Analysis 15, 124-139.

PORTER M. E. (1990) The Competitive Advantage of Nations. Macmillan, London..

QUIGLEY J. M. (1998) Urban Diversity and Economic Growth, Journal of Economic Perspectives 12 (2), 127-138.

ROSENTHAL S. S. and STRANGE W. C. (2004) Evidence on the Nature and Sources of Agglomeration Economies, in HENDERSON J. V. and THISSE J. F. (Eds) Handbook of Regional and Urban Economics, pp. 2119-2171. Elsevier, Amsterdam.

STORPER M. and VENABLES A. J. (2004) Buzz: face-to-face contact and the urban economy, Journal of Economic Geography 4 (2004), 351-370.

STORPER M. and WALKER R. (1989) The Capitalist Imperative. Territory, Technology and Industrial Growth. New York, Basil Blackwell:.

UTTERBACK J. M. and SUAREZ F. F. (1993) Innovation, Competition, and Industry Structure, Research Policy 22, 1-21.

Tables

Table 1: Externalities and their origins

| | Costs of production | Knowledge & skills | Other |
|---|---|---|---|
| Urbanization | land rents<br>wage premium<br>congestion | highly-skilled employees<br>knowledge infrastructure | market access |
| MAR | matching costs labour market | specialized labour force | access to specialized clients and suppliers |
| | minimize inventories | intra-industry knowledge spillovers | |
| | transportation costs within the value chain | joint innovation efforts within the value chain | |
| Jacobs' | low-risk environment | inter-industry knowledge spillovers | |
| | love of variety<br>lack of focus | | |



Table 2: Industry characteristics and life-cycle developments

| | Life cycle stage of industry | | |
|---|---|---|---|
| | Young | → | Mature |
| Innovation intensity | high | → | low |
| Type of innovation | product | → | process |
| Mode of competition | product | → | price |

Table 3: Agglomeration externalities and life-cycle dynamics

| | | | Life-cycle stage of industry | | |
|---|---|---|---|---|---|
| | | | Young | → | Mature |
| Urbanization | factor costs | high land rents | 0 | → | - |
| | | high wages | 0 | → | - |
| | | congestion | 0 | → | - |
| | knowledge | highly-skilled labour force | + | → | 0 |
| | | knowledge infrastructure | + | → | + |
| | market conditions | access to large market | 0 | → | + |
| | | access to sophisticated market | + | → | 0 |
| | | | | | |
| MAR | factor costs | low matching costs labour market | 0 | → | + |
| | | low inventories | 0 | → | + |
| | | low transportation costs within the value chain | 0 | → | + |
| | knowledge | large specialized labour force | 0 | → | + |
| | | high intra-industry knowledge spillovers | + | → | + |
| | | easy joint innovation efforts within the value chain | 0 | → | + |
| | market conditions | easy access to specialized clients and suppliers | 0 | → | + |
| | | | | | |
| Jacobs' | factor costs | large variety of services and goods | + | → | 0 |
| | | lack of focus | 0 | → | - |
| | knowledge | high inter-industry knowledge spillovers | + | → | 0 |
| | market conditions | reduced volatility in demand and supply | + | → | 0 |

+: expected effect is positive, 0: no important effects expected/effects cancelling out;-: negative effect expected.

Table 4: Definition of life-cycle stages

| | textiles | sawmilling | carpentry | furniture | paper | publishing | chemicals | other plastic | metalware | electric motors | communication | instruments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1974 | intermediate | intermediate | intermediate | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1975 | intermediate | intermediate | intermediate | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1976 | intermediate | intermediate | intermediate | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1977 | intermediate | intermediate | intermediate | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1978 | intermediate | intermediate | intermediate | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1979 | intermediate | mature | mature | intermediate | mature | intermediate | intermediate | young | young | mature | intermediate | young |
| 1980 | mature | mature | mature | intermediate | mature | intermediate | intermediate | intermediate | young | mature | intermediate | young |
| 1981 | mature | mature | mature | mature | mature | intermediate | intermediate | intermediate | young | mature | intermediate | young |
| 1982 | mature | mature | mature | mature | mature | intermediate | intermediate | intermediate | young | mature | intermediate | young |
| 1983 | mature | mature | mature | mature | mature | mature | intermediate | intermediate | young | mature | intermediate | young |
| 1984 | mature | mature | mature | mature | mature | mature | intermediate | young | young | young | mature | young |
| 1985 | mature | mature | mature | mature | mature | mature | intermediate | young | young | young | mature | young |
| 1986 | mature | mature | mature | mature | mature | intermediate | intermediate | young | young | young | mature | young |
| 1987 | mature | mature | mature | mature | mature | intermediate | intermediate | young | young | young | mature | young |
| 1988 | mature | mature | mature | intermediate | mature | intermediate | intermediate | young | young | young | mature | young |
| 1989 | mature | mature | mature | intermediate | mature | young | mature | intermediate | young | young | mature | young |
| 1990 | intermediate | mature | mature | intermediate | mature | young | mature | intermediate | young | young | intermediate | young |
| 1991 | intermediate | mature | mature | intermediate | mature | young | mature | intermediate | young | young | intermediate | young |
| 1992 | intermediate | mature | mature | intermediate | mature | young | mature | intermediate | young | young | intermediate | young |
| 1993 | intermediate | mature | mature | intermediate | mature | young | mature | intermediate | intermediate | young | intermediate | young |
| 1994 | intermediate | mature | mature | intermediate | mature | young | mature | intermediate | intermediate | young | intermediate | young |
| 1995 | mature | mature | mature | intermediate | mature | young | mature | intermediate | intermediate | young | young | young |
| 1996 | mature | mature | intermediate | mature | mature | young | mature | intermediate | intermediate | young | young | young |
| 1997 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 1998 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 1999 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 2000 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 2001 | mature | mature | intermediate | mature | mature | young | intermediate | mature | intermediate | young | young | young |
| 2002 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 2003 | mature | mature | intermediate | mature | mature | young | mature | mature | intermediate | young | young | young |
| 2004 | mature | mature | mature | mature | mature | intermediate | mature | mature | intermediate | young | young | young |

young
intermediate
mature

Table 5: General descriptive statistcs of variables.

| | mean | min | max | standard deviation between | within | ratio |
|---|---|---|---|---|---|---|
| log(VA) | 10.15 | -3.77 | 17.82 | 2.13 | 0.98 | 2.17 |
| log(L) | 4.55 | -9.36 | 10.04 | 1.97 | 0.73 | 2.71 |
| log(MAR) | 0.48 | -34.86 | 5.76 | 2.38 | 1.12 | 2.12 |
| log(JACOBS) | 0.75 | 0.00 | 3.33 | 0.67 | 0.26 | 2.54 |
| log(POPULATION) | 10.95 | 9.08 | 13.80 | 0.73 | 0.04 | 17.45 |
| log(HOUSE) | 6.09 | 4.56 | 8.31 | 0.28 | 0.50 | 0.55 |
| log(WAGE) | 0.01 | -0.60 | 0.26 | 0.04 | 0.03 | 1.29 |

Variables as defined in section 5. Final column contains ratio of between to within standard deviation.

Table 6: Correlation table of regressors.

| | log(L) | log(MAR) | log(JACOBS) | log(HOUSE) | log(WAGE) | log(POPULATION) |
|---|---|---|---|---|---|---|
| log(L) | 1.000 | | | | | |
| log(MAR) | 0.663 | 1.000 | | | | |
| log(JACOBS) | 0.283 | 0.351 | 1.000 | | | |
| log(HOUSE) | 0.054 | 0.062 | 0.168 | 1.000 | | |
| log(WAGE) | 0.123 | 0.176 | 0.408 | 0.082 | 1.000 | |
| log(POPULATION) | 0.263 | 0.323 | 0.764 | 0.359 | 0.641 | 1.000 |

Variables are as defined in section 5.

Table 7: Outcomes

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| log(L) | 0.999 *** (0.004) | 0.998 *** (0.004) | 1.015 *** (0.004) | 0.998 *** (0.002) | | |
| log(L_Y) | | | | | 1.008 *** (0.004) | 1.013 *** (0.004) |
| log(L_I) | | | | | 0.998 *** (0.004) | 0.998 *** (0.004) |
| log(L_M) | | | | | 0.985 *** (0.003) | 0.986 *** (0.003) |
| log(MAR) | 0.019 *** (0.004) | 0.019 *** (0.004) | 0.017 *** (0.004) | 0.019 *** (0.003) | | |
| log(MAR_Y) | | | | | 0.008 * (0.004) | 0.015 *** (0.004) |
| log(MAR_I) | | | | | 0.013 *** (0.005) | 0.021 *** (0.005) |
| log(MAR_M) | | | | | 0.022 *** (0.005) | 0.027 *** (0.005) |
| log(JACOBS) | 0.014 (0.011) | 0.014 (0.011) | 0.009 (0.010) | -0.017 *** (0.006) | | |
| log(JACOBS_Y) | | | | | 0.019 ** (0.009) | -0.004 (0.009) |
| log(JACOBS_I) | | | | | 0.006 (0.009) | 0.014 * (0.008) |
| log(JACOBS_M) | | | | | -0.053 *** (0.008) | -0.052 *** (0.008) |
| log(POPULATION) | 0.138 * (0.079) | 0.124 (0.090) | -0.002 (0.020) | 0.018 ** (0.009) | | |
| log(POPULATION_Y) | | | | | -0.029 *** (0.010) | 0.009 (0.008) |
| log(POPULATION_I) | | | | | 0.042 *** (0.010) | 0.007 (0.008) |
| log(POPULATION_M) | | | | | 0.038 *** (0.010) | 0.023 *** (0.008) |
| log(WAGE) | | 0.149 (0.102) | 0.144 (0.098) | 0.149 ** (0.072) | | |
| log(WAGE_Y) | | | | | 0.522 *** (0.114) | 0.341 *** (0.112) |
| log(WAGE I) | | | | | -0.136 (0.116) | 0.008 (0.113) |
| log(WAGE_M) | | | | | 0.074 (0.102) | 0.092 (0.100) |
| log(HOUSE) | | 0.000 (0.023) | 0.008 (0.020) | 0.000 (0.012) | | |
| log(HOUSE_Y) | | | | | 0.069 *** (0.014) | |
| log(HOUSE_I) | | | | | -0.063 *** (0.015) | |
| log(HOUSE_M) | | | | | -0.027 ** (0.014) | |
| Nobs | 17424 | 17424 | 17424 | 17424 | 17424 | 17424 |
| N | 753 | 753 | 753 | 753 | 753 | 753 |
| average T | 23.10 | 23.10 | 23.10 | 23.10 | 23.10 | 23.10 |
| Rsq | 0.873 | 0.873 | 0.873 | 0.970 | 0.970 | 0.970 |

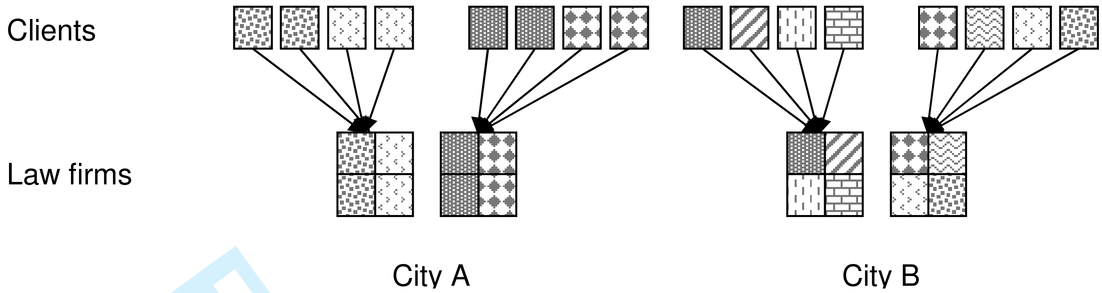
***: p<0.01, **: p<0.05, *: p<0.10. Dependent variable: log(VA). Variables as defined in section 5. _Y: young industries, _I: intermediate industries, _M: mature industries. All models include time and city-industry effects. Models in column 1 and 2 use fixed effects. Model in column 3 uses random effects. Models in column 4, 5 and 6 are based on FEVD estimators.

Table 8: Outcomes of alternative stage definitions

| | (5a) | (5b) | (5c) |
|---|---|---|---|
| log(L_Y) | 1.006 *** (0.004) | 1.008 *** (0.004) | 1.009 *** (0.006) |
| log(L_I) | 0.991 *** (0.006) | 0.998 *** (0.004) | 0.997 *** (0.003) |
| log(L_M) | 0.990 *** (0.003) | 0.985 *** (0.003) | 0.981 *** (0.005) |
| log(MAR_Y) | 0.007 * (0.004) | 0.008 * (0.004) | 0.003 (0.006) |
| log(MAR_I) | 0.025 *** (0.008) | 0.013 *** (0.005) | 0.019 *** (0.003) |
| log(MAR_M) | 0.021 *** (0.004) | 0.022 *** (0.005) | 0.051 *** (0.010) |
| log(JACOBS_Y) | 0.015 * (0.009) | 0.019 ** (0.009) | 0.026 * (0.014) |
| log(JACOBS_I) | 0.011 (0.013) | 0.006 (0.009) | -0.001 (0.006) |
| log(JACOBS_M) | -0.038 *** (0.007) | -0.053 *** (0.008) | -0.139 *** (0.013) |
| log(POPULATION_Y) | -0.010 (0.010) | -0.029 *** (0.010) | -0.059 *** (0.013) |
| log(POPULATION_I) | 0.039 *** (0.013) | 0.042 *** (0.010) | 0.024 *** (0.009) |
| log(POPULATION_M) | 0.037 *** (0.009) | 0.038 *** (0.010) | 0.058 *** (0.014) |
| log(WAGE_Y) | 0.429 *** (0.106) | 0.522 *** (0.114) | 0.613 *** (0.174) |
| log(WAGE_I) | -0.507 *** (0.187) | -0.136 (0.116) | 0.169 ** (0.081) |
| log(WAGE_M) | 0.078 (0.091) | 0.074 (0.102) | -0.445 *** (0.167) |
| log(HOUSE_Y) | 0.038 *** (0.014) | 0.069 *** (0.014) | 0.126 *** (0.021) |
| log(HOUSE_I) | -0.039 * (0.02) | -0.063 *** (0.015) | -0.017 (0.013) |
| log(HOUSE_M) | -0.033 ** (0.013) | -0.027 ** (0.014) | -0.031 (0.021) |

***: p<0.01, **: p<0.05, *: p<0.10. Dependent variable: log(VA). Variables as defined in section 5. _Y: young industries, _I: intermediate industries, _M: mature industries. All models include time and city-industry effects. Models use the FEVD estimator. (5a) young: $I_{it}^{norm}$ <1-0.1 std dev; intermediate: 1-.1 std dev < $I_{it}^{norm}$ <1+0.1 std dev; mature: $I_{it}^{norm}$ >1+0.1 std dev ; (5b) young: $I_{it}^{norm}$ <1-0.3 std dev; intermediate: 1-0.3 std dev < $I_{it}^{norm}$ <1+0.3 std dev; mature: $I_{it}^{norm}$ >1+0.3 std dev; (5c) young: $I_{it}^{norm}$ <1-0.6 std dev; intermediate: 1-.6 std dev < $I_{it}^{norm}$ <1+0.6 std dev; mature: $I_{it}^{norm}$ >1+0.6 std dev

Figure 1: Diversity and focus



[i] For a concise exposition on the different research traditions on urban economic growth and diversity, see QUIGLEY (1998). We focus on the empirical debate that began in the 1990s.

[ii] Another term for benefits of local specialization is "localization externalities". The difference is that MAR externalities are supposed to be dynamic or self-reinforcing, whereas localization externalities are more of a static nature. We do not think this distinction has much merit in this study, and will therefore call all benefits firms derive from local specialization MAR externalities.

[iii] A fourth type of externality arises from a contest between local producers for excellence and innovation (PORTER,1990). However, the intensity of competition is difficult to quantify. Therefore, it is not included here.

[iv] In the 1980s, the ILC approach was widely applied in economic geography to explain the rise of new industries in new growth regions (STORPER and WALKER 1989; BOSCHMA 1997). AUDRETSCH and FELDMAN (1996) found evidence that the propensity for innovative efforts to cluster spatially was shaped by the characteristics of life-cycle phases. Also in GLAESER et al. (1992), mention is made of interactions between the maturity of an industry and the agglomeration externalities it may experience. However, the authors' suspicion that agglomeration externalities – regardless of type – are strongest in young industries is only mentioned in passing and as a possible caveat that should be kept in mind when interpreting their results.

[v] The data on which the databases are built were provided by Statistics Sweden. We have run several algorithm-based data cleaning procedures and parts of the data were checked by hand. Information can be obtained from the authors.

[vi] This was the case for chemicals, furniture, metal ware, paper, publishing, and textiles.

[vii] We choose an exponential decay that gives a 1% weighting to places 100 km away. Within a circle of 10 km surrounding the A-region-core, decay was assumed to be zero. We also experimented with distance decays that gave a 1% weighting to places 75 or 125 km away. Results, available on request, were very similar to the ones presented in this text.

viii The data on housing prices are based on sales prices for smaller houses, available at the municipality level. Before 1981, only growth rates in housing prices at the province level and for major metropolitan areas (11 regions in total) were available. Between 1974 and 1981, our housing prices are therefore only estimates.

ix For example, our smallest industry, electric motors, consists (on average) of only 111.6 plants. Moreover, for over one third of all years, net entry into single industries is between just -5 and +5 plants.

x As the threshold is arbitrary, it is important to test the sensitivity of results to changes in this threshold. We come back to this issue in section 6.

xi In the analyses below, we use an unweighted average of plant wages to calculate $w_{cit}$ and $\overline{w}_{it}$. As a robustness check, we also used weighted averages that use plant employment levels as weights. The outcomes are all but indistinguishable from the ones we present in this article.

xii The effect of the number of plants while holding local employment constant can also be interpreted as the effect of average plant size. However, average plant size is not very informative in our study, as plant sizes are very skewed. Therefore, we do not attribute any specific meaning to this interpretation.

xiii E.g., the HHIs for the employment sets {100,20,20,20} and {100,100,9,9} are about the same.

xiv We set this threshold at 10 plants, but also experimented with different threshold values and variants based on employment. This had virtually no impact on the outcomes.

xv Different lag structures have been tested and yielded very similar outcomes.

xvi In the first step, they estimate an FE model. The residuals of this equation now contain two components: the unobserved city-industry effects and a part that can be explained using variables with no or very little variation over time. In the next step, the authors regress these residuals on the time-invariant and slowly-changing variables, and then decompose them into two parts: an unexplained part and a part explained by the time-invariant and slowly-changing variables. In the final step, the complete model is rerun without fixed effects, but this time with estimates of the unexplained part of the city-region effects obtained in the second step.

[xvii] This is a common problem. In many studies, the only available instruments are weak and yield such low efficiency that no conclusions can be drawn from the analyses (e.g., HENDERSON, 2003).

[xviii] Remember that due to the time-varying specification of MAR, the estimate depends only on within-variation. Cross-sectional differences do not have any influence.